nternational Symposium (2009)
Rice Research in the Era of Global Warming
43~55

Marker-Assisted Breeding for Abiotic Stress Tolerance in Rice: Progress and Future Perspectives

Michael J. Thomson[1,*]

[1]International Rice Research Institute (IRRI), DAPO Box 7777, Metro Manila, Philippines
*Corresponding author:m.thomson@cgiar.org

ABSTRACT

While global climate change threatens to increase the occurrence and intensity of production constraints such as abiotic stresses, new marker-assisted breeding techniques may enable the rapid development of stress-tolerant rice varieties to face these challenges. Through the mapping of quantitative trait loci (QTLs) and precise transfer of tolerant QTLs by marker-assisted backcrossing (MABC), improved varieties can be developed with beneficial traits. For example, salt stress is a major constraint across many rice-producing areas because of the high sensitivity of modern rice varieties. At IRRI, we are working towards combining QTLs conferring salt tolerance to develop robust cultivars for stress-prone environments. Previously, RILs from a cross between IR29 and an accession of the tolerant landrace Pokkali were used to map several salt tolerance QTLs, including *Saltol*, a major QTL on the short arm of chromosome 1. Near-isogenic lines were developed between IR29 and Pokkali, and subsequently multiple Pokkali alleles were identified between the original RILs and the NILs. The Pokkali introgressions in six advanced lines were also characterized using an Illumina GoldenGate assay developed at Cornell University to genotype 1,536 SNP loci across the rice genome. A precision marker-assisted backcrossing

system is also being employed to efficiently transfer beneficial Pokkali introgressions into popular varieties such as BRRI dhan28, an important dry season variety in coastal Bangladesh, to increase production in saline prone areas. Future advances in new marker-assisted breeding strategies, rapid sequencing of multiple rice genomes, and SNP genotyping technologies also promise to increase the efficiency of marker development and integration of molecular tools into modern breeding programs.
Key words: *Orzya sativa*, Molecular breeding, Salinity tolerance, SNP genotyping.

INTRODUCTION

Although plant breeders have made significant progress using conventional breeding methods, the past 20 years have seen the potential of molecular markers gain prominence in breeding programs. Even though the types of molecular markers have changed over the years, the basic concept is the same: markers allow tracking of polymorphisms based on underlying changes in the DNA of each individual. These markers can then provide information on genetic relatedness, linkage to important traits, and detection of donor introgressions in segregating populations. Ultimately, the usefulness of integrating molecular markers into a breeding program will depend on the how well the markers can replace phenotyping, both in terms of cost-effectiveness and in their ability to predict performance under different environments. Although great efforts have been invested in mapping genes and in developing marker techniques, only recently have markers become routinely used in major breeding programs, as evidenced by the successful release of the first generation of marker-assisted breeding products from both the public and private sector. There are a number of strategies now available to incorporate molecular tools into a rice breeding program to increase the efficiency and power of selection, as described below.

Mapping Quantitative Trait Loci (QTLs)

While early success was made with mapping genes for qualitative traits, many traits in rice breeding are quantitative traits showing a normal distribution, due to control by multiple genes and environmental interactions. Mapping QTLs, which are chromosomal regions that contain a gene or genes contributing to a trait phenotype, requires comprehensive genetic maps to assign a phenotypic effect at markers near the QTL. Numerous QTL studies have led to an explosion in the number of QTLs mapped in rice for a multitude of traits, with over 8,000 QTLs now available in the online Gramene database (www.gramene.org). While there are dozens of QTLs mapped for each major trait, the successful integration of QTLs into MAS programs has been more elusive, likely due to the fact that many traits are controlled by multiple small effect QTLs, often complicated by environmental interactions and/or differing genetic background effects. Care must be taken to select the most promising QTL targets for breeding applications, focusing on large effect QTLs that are stable across environments and genetic backgrounds.

Exploring Novel Genetic Variation

The basis for future improvement in breeding programs is the successful identification and use of novel genetic variation. Much work has been done to characterize the genetic diversity across rice varieties on a global scale, but more research is needed to investigate diversity in more detail on the national and local scales. For example, the genetic diversity of traditional rice varieties in Indonesia revealed important differences between representative rice accessions across a geographically diverse sample versus newly collected varieties in an isolated region in Borneo (Thomson et al. 2007; Thomson et al. 2009). While it can be difficult to identify and extract useful variation directly from germplasm collections, one advantage of QTL mapping is that it can enable the discovery of

hidden diversity in exotic germplasm that is not readily accessible to breeders (McCouch et al. 2007). For example, a QTL study between a U.S. variety and an accession of *Orzya rufipogon* identified transgressive QTLs from the wild rice relative that could improve the elite culitvar (Thomson et al. 2003). A large investment is often required to further characterize target QTLs through near-isogenic line development and subsequent fine-mapping, as was performed for *O. rufipogon*-derived QTLs for grain weight, heading date, and red pericarp (Li et al. 2004; Thomson et al. 2006; Sweeney et al. 2006).

Marker-Assisted Backcrossing (MABC)

One of the simplest forms of marker-assisted breeding is the use of molecular markers to improve the conventional backcross conversion method, where the desired trait is transferred to another line through crossing followed by repeated backcrossing to the recurrent parent to reconstitute the original variety. The use of markers can speed up the conversion process and greatly reduce the amount of negative linkage drag correlated with the target locus. Consequently, the technique of marker-assisted backcrossing (MABC) has been promoted as a way to take beneficial QTLs and incorporate them into breeding lines using foreground markers at the target locus and background markers across the rest of the genome. Thus the new paradigm for non-transgenic molecular breeding follows a pathway from identifying landraces as donors for the trait of interest, to QTL discovery and fine-mapping of the most promising loci providing tolerance, and finally using marker-assisted backcrossing to rapidly transfer the QTLs to popular varieties. Recently, the power and efficacy of this approach was proven through the discovery, fine-mapping and cloning the submergence tolerance QTL *Sub1* from the tolerant landrace FR13A, followed by successful MABC transfer of *Sub1* into the popular variety Swarna and five other mega-varieties (Xu et al. 2006; Neeraja et al. 2007; Septiningsih et al. 2009). In practice, the efficiency

of a MABC program depends on a number of factors, including the availability of markers in the target region, the precision of the target gene/QTL fine-map, the rate of polymorphism when identifying background markers, and the cost, speed, and failure rate of the markers employed in each customized MABC system.

Pyramiding Genes and QTLs

The use of molecular markers can also enable breeders to pyramid, or combine, multiple genes or QTLs into the same variety. In some cases, such as the pyramiding of resistance loci conferring resistance to a disease, relying on phenotype alone would be impossible, but using markers allows the tracking and combination of specific resistance genes. Pyramiding can also be used to "stack" genes for multiple traits into the same variety. For example, gene pyramiding has been successfully used in rice to combine multiple genes for bacterial leaf blight, as well as resistance genes for other pests and diseases (for review, see Collard and Mackill, 2008). When dealing with quantitative traits, such as abiotic stress tolerance, QTL pyramiding as been proposed as a strategy to combine multiple QTLs, each controlling a trait by a different mechanism, to increase the contribution to the trait effect and for developing optimal combinations of QTLs for multiple traits.

Gene Discovery and Functional Markers

Although foreground selection using closely linked markers for major genes and QTLs is the most common approach for marker-assisted breeding, a refinement of this strategy is to target the genes themselves for marker development. Intense research efforts around the world are focused on dissecting the molecular mechanisms and genetic pathways underlying key traits in rice—leveraging the value of the complete genome sequence for gene discovery and functional analysis. One common justification of the massive research investment in rice

functional genomics is that the end result will empower more efficient strategies for molecular breeding. Thus genomics-assisted breeding promises to apply discoveries from mutant analysis, DNA microarrays, QTL cloning and allele mining to develop more precise molecular tools for crop improvement (Varshney et al. 2005; Leung 2008). The most intuitive way to do this is to target a functional nucleotide polymorphism, which is a change in a gene that causes the desired phenotype, for development of functional markers (Andersen and Lubberstedt, 2003). By relying directly on the causal polymorphism, these "perfect" markers will be diagnostic of the favorable allele, since they will always co-segregate with the trait phenotype. For example, once the gene controlling aroma in rice was cloned, the sequence polymorphisms that led to increased aroma were identified and developed into a perfect marker for fragrance in rice (Bradbury et al. 2005). As more genes controlling key traits in rice are characterized, there will be more opportunities to mine superior alleles from germplasm collections and to develop functional markers for more precise and efficient marker-assisted selection.

RESULTS AND DISCUSSION

Breeding for Salinity Tolerance in Rice

Salinity is a problem in coastal regions due to intrusion of brackish water during the dry season and a worsening problem in inland areas with saline soil and/or poor quality irrigation water. There are a number of traditional varieties adapted to these regions which show a remarkable tolerance to salinity and have been used to study the basis of tolerance in rice. Tolerance to salinity may involve a number of different physiological mechanisms, including sodium exclusion from roots, controlled sodium transport between root and shoot, sequestering of sodium in older tissues and in the vacuoles, better stomatal control, and antioxidant activity (Ismail et al. 2007). As traditional landraces have shown tolerance

through different mechanisms, a pyramiding approach has been proposed to combine complementary mechanisms to develop rice varieties with improved levels of tolerance. Progress has been made to transfer salinity tolerance from traditional donors to modern high yielding varieties, but the complexity of the trait has proven challenging to breeders. Since most of the tolerant landraces are low-yielding varieties with numerous undesirable traits, a precision marker-assisted breeding approach is needed to rapidly transfer the high levels of tolerance of the landraces into high yielding breeding lines while at the same time avoiding negative linkage drag.

QTL Mapping and MABC for Salinity Tolerance

Previously, a recombinant inbred line (RIL) mapping population between IR29 and the salt tolerant landrace Pokkali was used to identify a major QTL for seedling-stage salinity tolerance, named *Saltol*, on the short arm of chromosome 1, along with a number of minor QTLs. Subsequently, RFLP and SSR markers were added to the *Saltol* region to more precisely define the position (Bonilla et al. 2002). Near-isogenic lines were developed between IR29 and Pokkali, and multiple Pokkali alleles were identified between the original RILs and the NILs. The Pokkali introgressions in six advanced lines were also characterized using an Illumina GoldenGate assay developed by Susan McCouch at Cornell University to genotype 1,536 SNP loci across the rice genome. Other Pokkali QTLs are being targeted for NIL development to allow testing of individual QTL effects. Furthermore, mapping populations using additional sources of tolerance are also being developed at IRRI to compare QTLs across different tolerant donors under both seedling and reproductive stages. A precision marker-assisted backcrossing system is also being employed to efficiently transfer beneficial Pokkali introgressions into popular varieties such as BRRI dhan28, an important dry season variety in coastal Bangladesh. These new salt tolerant varieties can then

help increase production in stress-prone environments, which will likely pose increasingly difficult growing conditions in the future.

CONCLUSIONS

Future Perspectives

Advances in genomics promise to enable more powerful trait dissection through techniques such as global genome expression analysis, rapid re-sequencing of whole genomes, and association studies using diverse germplasm (Zhu et al. 2008). This wealth of data can be combined with QTL analysis using a positional candidate gene approach to identify converging lines of evidence pointing towards the key genetic players that control abiotic stress tolerance pathways. Moreover, knowledge of the genetic pathways can be integrated with a detailed physiological analysis of tolerance mechanisms to achieve a more complete picture of the rice plant's response to salinity at the whole plant level. Although QTL discovery, fine-mapping, and marker-assisted backcrossing is a straightforward and powerful technique, there are still many challenges to be faced when applying this approach to different abiotic stresses. For example, the QTL effect itself may vary depending on the genetic background, environmental conditions, and plant growth stages. To offset these problems, only major QTLs of large effect should be pursued, as these are more likely to be stable and can justify the substantial investment required for fine-mapping and MABC.

New Marker-Assisted Breeding Strategies

Recent strategies have proposed selection approaches based on increasing the frequency of desirable alleles for larger numbers of QTLs in a population, for example up to 30 loci using marker-assisted recurrent selection (MARS; Bernardo, 2008). In MARS, multiple cycles of marker-assisted selection are

performed: F2 individuals with the desirable QTL alleles are crossed, and the progeny are screened with markers to undergo another round of selection, which continues until the desired frequency of favorable QTL alleles is achieved. In practice, the markers are used to calculate a genotypic value for each individual, and crosses are made based on this value to increase the frequency of favorable alleles in the population, which then reduces the size of the population required to identify an ideal genotype (Eathington et al. 2007). MARS schemes have proven popular with commercial breeding programs, and recently there has been more interest to use this strategy in the public sector as well, such as with the Molecular Breeding Platform proposed by the Generation Challenge Program and the Gates Foundation. Another strategy, referred to as genome-wide (or genomic) selection, proposes to use the combined effect on a trait from all the markers covering the genome instead of just using the significant loci as in MARS (Bernardo 2008). In this case, both phenotyping and genotyping the population is still needed to calculate the genotypic values of the progeny before selection, but information on specific QTLs is not necessary, since all markers are taken into account. Furthermore, efforts are being made to focus on predicting the performance of progeny from selected crosses even before phenotyping, which becomes possible once a large database of marker-trait associations is developed, as is currently being used in the Monsanto maize breeding program (Eathington et al. 2007). If this system will be employed in rice, it will require genome-wide marker scans to be performed on breeding lines that are then phenotyped in multi-environment field trials.

Sequencing More Rice Genomes

The completion of the high-quality DNA sequence of the rice genome by the International Rice Genome Sequencing Project was a landmark achievement. Although it took over eight years and was a massive undertaking, the fact remains

that it was a single rice genome, of the temperate *japonica* variety Nipponbare, that was sequenced out of the thousands of different rice genomes existing in the diverse germplasm from around the world. Nonetheless, there is tremendous value in having a "gold standard" reference sequence available, which will enable a host of downstream applications, not the least of which will be the re-sequencing of many more rice genomes in the years to come using the high-quality sequence of Nipponbare as a reference to align sequence fragments and provide added value to the 93-11 sequence data and other genomes to follow. At the same time, it will be important to have a *de novo* high quality *indica* and *aus* genome sequences as well, since there may be unique portions of the genome in those subgroups that are not present in Nipponbare, that would be hidden from any sequence assembly depending on the Nipponbare sequence, as was seen in a recent study on the *Pup1* QTL region which had several genes unique to Kasalath (Heuer et al. 2009). With cost-effective and rapid 2nd-generation sequencing techniques (such as Illumina Genome Analyzer IIx) whole genome sequencing of rice varieties will soon generate massive amounts of sequence data that will provide a valuable resource to the rice community.

SNP Genotyping Technologies

The successful implementation of these marker-assisted breeding strategies is dependent on having an efficient and robust genotyping system in place (Collard et al. 2008). For many years, simple sequence repeats (SSRs) have been the marker system of choice due to high polymorphism rates and the ability to run them on inexpensive gel electrophoresis equipment found in most labs. However, the routine integration of markers into modern breeding programs will require high-throughput genotyping platforms that can handle large numbers of samples at a low cost. Thus a new generation of markers based on single nucleotide polymorphisms (SNPs) is now rapidly overtaking SSRs due to new

SNP genotyping platforms that offer multiplexed sets of markers for different applications. SNP discovery efforts have begun to make large pools of SNPs available to the rice community (McNally et al. 2009; www.oryzasnp.org). Efforts are now being made to develop functional SNPs for foreground selection, to optimize 384 or 1536-plex SNP chips for low resolution genome scans, and to develop high resolution SNP chips with >40,000 SNPs for association genetics studies (personal communication, Susan McCouch, Cornell University). For example, an Illumina BeadXpress Reader at IRRI can genotype 384-SNPs on 96 samples with a two day protocol at less than $0.10/data point. As these genotyping tools are made more accessible to rice breeders, the full potential of marker-assisted breeding in mainstream breeding programs will finally be realized.

ACKNOWLEDGMENTS

The work was funded in part through the Challenge Program on Water and Food, the German Federal Ministry for Economic Cooperation and Development (BMZ), the CGIAR Generation Challenge Program, the Bill and Melinda Gates Foundation, and the International Rice Research Institute.

REFERENCES

Andersen, J. and T. Lübberstedt. 2003. Functional markers in plants. Trends Plant Sci. 8:554-560.

Bernardo R. 2008. Molecular markers and selection for complex traits in plants: learning from the last 20 years. Crop Sci. 48:1649-1664.

Bonilla P., J. Dvorak, D. Mackill, K. Deal and G. Gregorio. 2002. RFLP and SSLP mapping of salinity tolerance genes in chromosome 1 of rice *Oryza sativa* L. using recombinant inbred lines. Philipp. Agric. Sci. 85:68-76.

Bradbury L., R. Henry, Q. Jin, R. Reinke and D. Waters. 2005. A perfect marker for fragrance genotyping in rice. Mol. Breeding 16:279-283.

Collard B., C. Vera Cruz, K. McNally, P. Virk and D. Mackill. 2008. Rice molecular breeding laboratories in the genomics era: current status and future

considerations. Int. J. Plant Genomics Volume 2008, Article ID 524847, doi:10.1155/2008/524847.

Collard B. and D. Mackill. 2008. Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. Phil. Trans. R. Soc. B 363:557-572.

Eathington S., T. Crosbie, M. Edwards, R. Reiter and J. Bull. 2008. Molecular markers in a commercial breeding program. Crop Sci. 47:S154-S163.

Heuer S., X. Lu, J.-H. Chin, J. Tanaka, H. Kanamori, T. Matsumoto, T. De Leon, V. Ulat, A. Ismail and M. Wissuwa. 2009. Comparative sequence analysis of the major quantitative trait locus phosphorus uptake 1 *Pup1* reveal a complex genetic structure. Plant Biotech. J. 7:456-471.

Ismail A., S. Heuer, M. Thomson and M. Wissuwa. 2007. Genetic and genomic approaches to develop rice germplasm for problem soils. Plant Mol. Biol. 65:547-570.

Jena K. and D. Mackill. 2008. Molecular markers and their use in marker-assisted selection in rice. Crop Sci. 48:1266-1276.

Leung H. 2008. Stressed genomics-bringing relief to rice fields. Curr. Opin. Plant Biol. 11:201-208.

Li J., M. Thomson and S. McCouch. 2004. Fine mapping of a grain weight QTL in the peri-centromeric region of rice chromosome 3. Genetics 168:2187-2195.

McCouch S., M. Sweeney, J. Li, H. Jiang, M. Thomson, E. Septiningsih, J. Edwards, P. Moncada, J. Xiao, A. Garris, T. Tai, C. Martinez, J. Tohme, M. Sugiono, A. McClung, L. Yuan and S. Ahn. 2007. Through the genetic bottleneck: *O. rufipogon* as a source of trait-enhancing alleles for *O. sativa*. Euphytica 154:317-339.

McNally K., K. Childs, R. Bohnert, R. Davidson, K. Zhao, V. Ulat, G. Zeller, R. Clark, D. Hoen, T. Bureau, R. Stokowski, D. Ballinger, K. Frazer, D. Cox, B. Padhukasahasram, C. Bustamante, D. Weigel, D. Mackill, R. Bruskiewich, G. Rätsch, C. Buell, H. Leung, and J. Leach. 2009. Genomewide SNP variation reveals relationships among landraces and modern varieties of rice. Proc. Natl. Acad. Sci. USA 106:12273-12278.

Neeraja C., R. Maghirang-Rodriguez, A. Pamplona, S. Heuer, B. Collard, E. Septiningsih, G. Vergara, D. Sanchez, K. Xu, A. Ismail and D. Mackill. 2007. A marker-assisted backcross approach for developing submergence-tolerant rice cultivars. Theor. Appl. Genet. 115:767-776.

Septiningsih E., A. Pamplona, D. Sanchez, R. Maghirang-Rodriguez, C. Neeraja, G. Vergara, S. Heuer, A. Ismail and D. Mackill. 2009. Development of Submergence-tolerant rice cultivars: The *Sub1* locus and beyond. Ann. Bot. 103:151-160.

Sweeney M., M. Thomson, B. Pfeil and S. McCouch. 2006. Caught red-handed: Rc encodes a basic helix-loop-helix protein conditioning red pericarp in rice. Plant Cell 18:283-294.

Thomson M., T. Tai, A. McClung, X.-H. Lai, M. Hinga, K. Lobos, Y. Xu, C. Martinez and S. McCouch. 2003. Mapping quantitative trait loci for yield, yield components and morphological traits in an advanced backcross population between *Oryza rufipogon* and the *Oryza sativa* cultivar Jefferson. Theor. Appl. Genet. 107:479-493.

Thomson M., J. Edwards, E. Septiningsih, S. Harrington and S. McCouch. 2006. Substitution mapping of *dth1.1*, a flowering time quantitative trait locus QTL associated with transgressive variation in rice, reveals multiple sub-QTL. Genetics 172:2501-2514.

Thomson M., E. Septiningsih, F. Suwardjo, T. Santoso, T. Silitonga and S. McCouch. 2007. Genetic diversity analysis of traditional and improved Indonesian rice *Oryza sativa* L. germplasm using microsatellite markers. Theor. Appl. Genet. 114:559-568.

Thomson M., N. Polato, J. Prasetiyono, K. Trijatmiko, T. Silitonga and S. McCouch. 2009. Genetic diversity of isolated populations of Indonesian landraces of rice *Oryza sativa* L. collected in East Kalimantan on the island of Borneo. Rice 2:80-92.

Varshney R., A. Graner and M. Sorrells. 2005. Genomics-assisted breeding for crop improvement. Trends Plant. Sci. 10:621-630.

Xu K., X. Xu, T. Fukao, P. Canlas, R. Maghirang-Rodriguez, S. Heuer, A. Ismail, J. Bailey-Serres, P. Ronald and D. Mackill. 2006. Sub1A is an ethylene-response-factor-like gene that confers submergence tolerance to rice. Nature 442:705-708.

Zhu C., M. Gore, E. Buckler and J. Yu. 2008. Status and prospects of association mapping in plants. The Plant Genome 1:5-20.